17. Stochastic Modeling of Human Learning Behavior

Jacob L. Meiry
Massachusetts Institute of Technology

The application of a stochastic model of human learning behavior in manual control tasks is extended to the performance of the human operator in regulating all second-order dynamic processes. In particular, human control decisions for optimal multiswitch and unstable dynamic systems are found to correlate well with responses recorded by the model in the study of simulated subjects. The universality of the model is discussed in the light of a series of compensatory task experiments.

INTRODUCTION

A theory of human learning behavior in the regulation of a second-order dynamic process by actuating a two position controller was previously presented in a paper entitled "Stochastic Modeling of Human Learning Behavior " in reference 1. The concept of skill organization is outlined as a statistical revision process applied to the identification of a control policy for the manual control task. To characterize the decisionmaking of the human operator, we have postulated a single-channel information system for the subject. A hypothetical decision center is entrusted with selecting among response alternatives at the moment of choice. The theory represents individual preferences for control policies as probabilities, continuously revised on the basis of experimental outcomes. Specifically, our interpretation of Bayesian statistics is the principal factor underlying the process of human decisionmaking.

The model of human learning behavior was submitted as a uniform control strategy applicable to all second-order, first-degree dynamic processes. Its success in representing the time-variant random nature of human learning in a series of controlled experiments was unqualified. This paper reports a new sequel of experiments to verify our theory and model operation when called to control harmonic oscillators and unstable plant dynamics. Finally, learning of a compensatory task is compared with human operator performance in controlling a pure inertia plant to minimize the integrated absolute error.

THEORY REVIEW

The stochastic information system postulated for the human operator consists of a sensor, a decision center, and an effecter, all serially connected. The sensor can perceive quantized state information since we assume a finite grid to overlay the state space of the

dynamic process. The decision center in turn processes the state information and decides upon a response. For the ideal relay controller this involves a decision: whether or not control polarity should be reversed. To visualize the process, the concept of a switch curve is helpful. By definition, the switch curve is the locus of all points in the state plane where a control polarity reversal should occur to force the second-order process' trajectory through the null state. As a matter of control policy, the human operator has to decide whether the state of the process has reached the location of the switch curve. We have proposed the use of subjective probabilities for this identification procedure. Thus we assume that stored in the decision center memory are columns of probabilities expressing the subjective estimate of the mesh where the switch curve lies, with higher probability reflecting higher confidence in the location of the curve. The strategy of the control policy was made dependent upon a switching probability evaluating the belief that the switch curve lies beneath the present state of the process. Indeed, if the switching probability is high, one is quite confident that a switch should take place. Selection of alternatives, then, can be based on the value of this probability. However, to account for variations in human control behavior, it is assumed that the decision to switch is a Bernoulli trial with the switching probability being the probability of success.

It is quite evident that learning a psychomotor task, then, is a revision of opinion about the location of the switch curve in the state space. In terms of probabilities, a weighting of previous decisions could be used to improve one's concepts about the appropriate strategy to follow. This is weighting of evidence, used here to evaluate the significance of an outcome of a decision in relation to the expected outcome. The revision of prior probabilities was assumed to constitute an application of Bayesian statistics, although this efficient algorithm exaggerates the decisionmaking capabilities of the human. Our theory as presented previously examines the information used as evidence and the formulation of posterior probabilities for future trials. The reader is referred to the above-mentioned paper for the mathematical treatment of the theory and the model.

THE EXPERIMENT

Three groups of eight subjects at MIT's man-vehicle laboratory were given the opportunity to learn a manual control task. They were briefed on the task and familiarized with the apparatus but were not allowed to practice prior to the first trial. The subject was required to null the initial misalinement between two line segments displayed on an oscilloscope by actuating a two position switch. The left segment remained stationary and the displacement of the other relative to it, x, satisfied a second-order differential equation. Each subject was given 50 5-second trials starting with the same initial conditions at 10-second intervals. Subject performance on each trial was measured by computing the integral of the absolute value of x over the 5 seconds, and this score was reported to the subject immediately after each trial. The dynamic processes controlled were of the form

$$\ddot{x} + x = u \qquad u = 4 \text{ cm/sec}^2$$

$$\ddot{x} + 4x = u \qquad u = 7.5 \text{ cm/sec}^2$$

$$\ddot{x} - |x| = u \qquad u = 7.5 \text{ cm/sec}^2$$

to represent the range of distinct response characteristics of second-order dynamics. A fourth group was engaged in learning to control a compensatory tracking task with the sum of 10 sine waves as input. The input signal is in accordance with that of reference 2. The process dynamics were pure inertia ($\ddot{x} = u$).

RESULTS AND DISCUSSION

The first series of experiments dealing with a regulator task is quite evidently designed to answer the following questions:

(1) Are model and subjects learning the optimal control strategy for difficult dynamic processes?

(2) Are model and subject responses drawn from the same parent population?

(3) Is a unique model of control strategy representative of decisionmaking in control of second-order dynamic processes?

To underline the significance of these experiments in evaluating the theory of human learning behavior, let us briefly discuss the plant dynamics selected. Since the optimal system response (with respect to the minimum integrated absolute error) is practically identical to the time optimal system (ref. 3), an easy comparison of control quality is feasible if time optimal control is used as a reference. Unlike the pure inertia dynamic system, the harmonic oscillator has imaginary eigenvalues. The time optimal control of the oscillator can involve multiple switching unlike a real-root eigenvalue system. Moreover, the optimal trajectories are a series of semicircles forcing the human operator to assume different control strategies for different parameters of the oscillator. One would expect the subjects to adopt a suboptimal policy whenever more than one switch is required. Thus the simulated dynamics of two oscillators were examined with expectations of finding a suboptimal control behavior for the oscillator with $\omega = 2$ rad/sec. On the other hand, the unstable plant is characterized by a switch curve with a very large gradient, while the harmonic oscillator has a relatively flat switch curve. Quite obviously the learning potential of the model and the subjects are heavily taxed with the dynamic processes selected, offering an opportunity for strict examination of the validity of the human learning behavior model.

EXPERIMENTAL RESULTS

We chose to express the subjects' learning performance by location of the switching trajectory in the phase plane. Figure 1 represents the average location of the first switching point measured for eight subjects at their 50th trial. Similarly the switch curve of the model is obtained from the table of posterior probabilities for eight simulated subjects. Figures 2, 3, and 4 are the average learning curves in terms of the integrated absolute error (IAE) for subjects and model. While all learning curves converge to an asymptotic value, some differences are noted between respective curves for subjects and model. We submit that these differences are not significant statistically, and we support this finding with results from the Mann-Whitney U test (ref. 4) performed on the interresponse time (time between two consecutive responses). At the 1 percent level, our tests accept the hypothesis that the subject sample and the model sample are drawn from identical distributions for 95 percent of the cases we examined in this manner (ref. 5).

Our statistical tests supporting the assumption of statistical images for a sample of behavioral simulations obtained with the model and from a series of psychomotor experiments offer us a quantitative basis to conclude that the proposed theory passes the test of generality. Similarly, it is quite evident (see fig. 1) that the skilled control policies of subjects and model closely resemble each other and are executed in a nearly time optimal fashion. All factors considered, we accept the theory of human learning behavior as an explanation of human control behavior in the regulation of second-order dynamic processes simulated as manual control tasks.

The application of the theory to compensatory tracking tasks is not an obvious extension of the class of problems just treated. First, it is evident that for such tasks the evidence for the revision process is heavily influenced by the dynamic characteristics of the input. Furthermore, an ideal two level relay controller for a manual compensatory task is quite unrealistic as a design from the point of view of fuel management. On the other hand, a study of model behavior in compensatory tasks is indicative of possible human difficulties in the search process for correct control policy. We submit that the original model of human learning behavior is capable of identifying a region of switching in the phase plane. After 20 trials, this region will contain over 90 percent of the switch reversals of the human operator in the compensatory task. Although these findings might support a speculation about the universality of the model as a representation of the time varying random performance of the human operator in manual control tasks, we surmise that only the approach of the theory is general. In this context, we have offered sufficient evidence drawn from a variety of tasks to support a contention about accepting the theory's explanation of human learning.

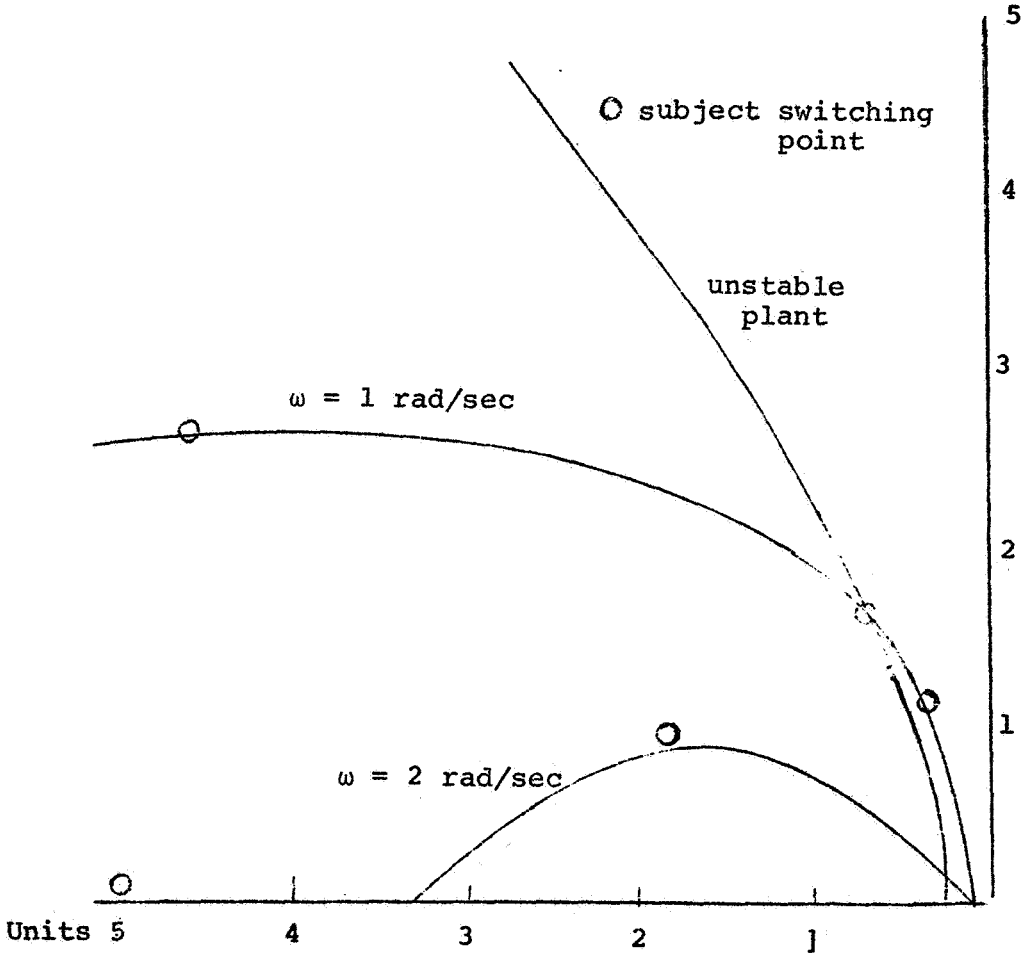
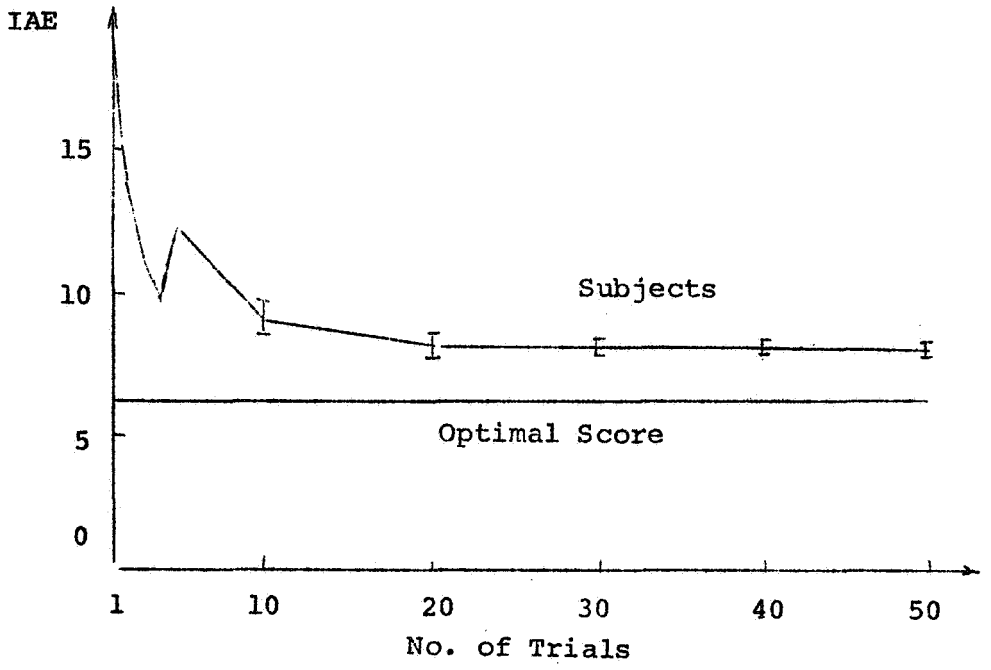


Figure 1.—Decision switching curves.



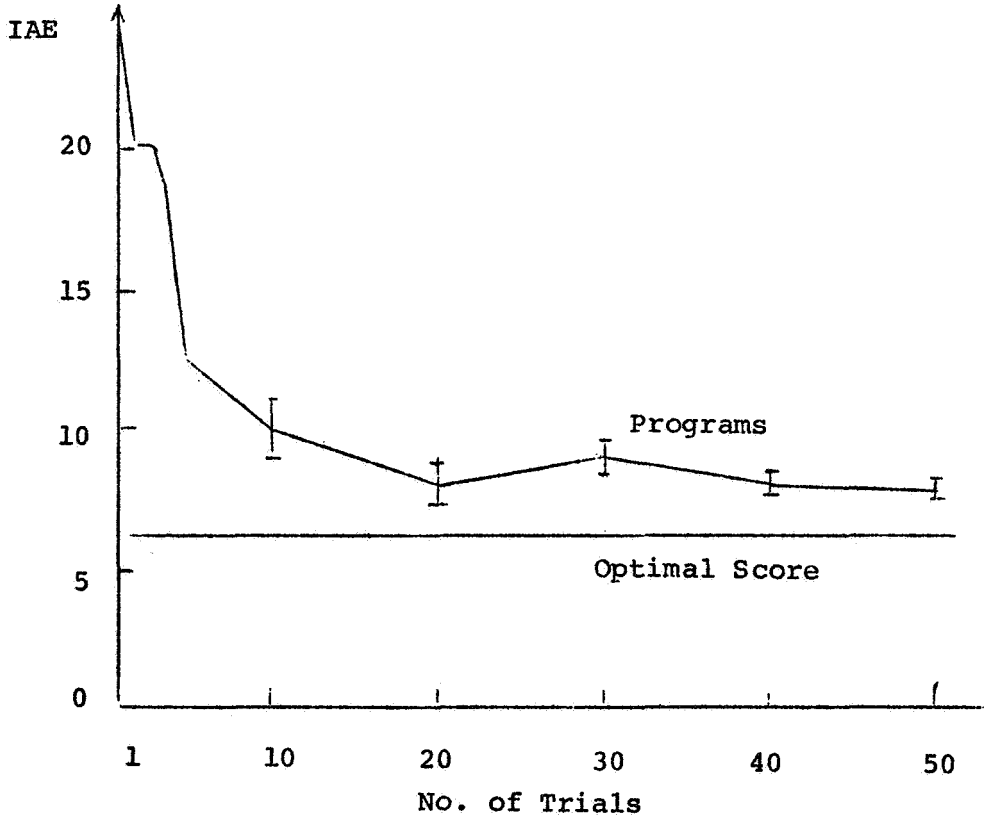


Figure 2.—Unstable plant learning curves.

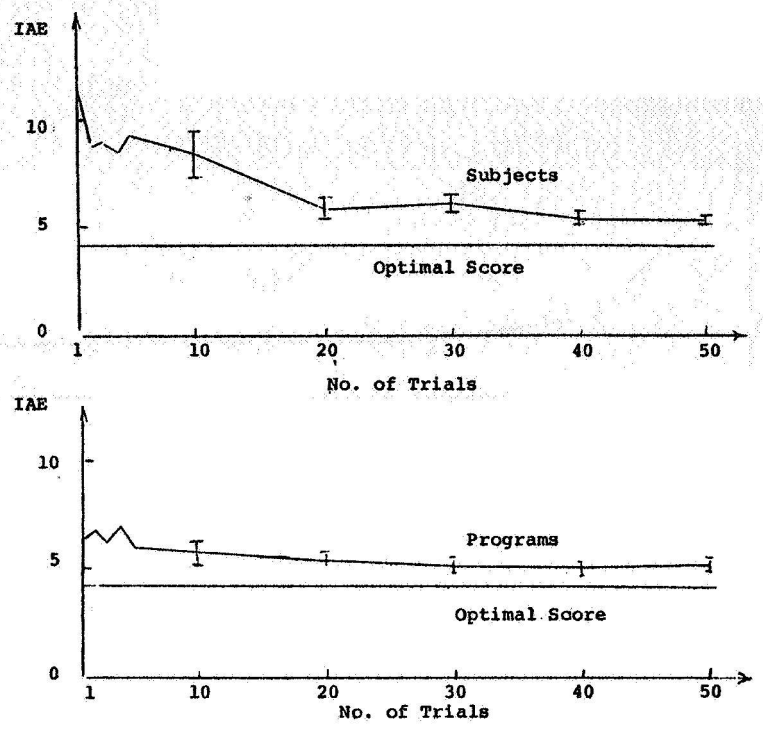


Figure 3.—Harmonic oscillator learning curves ($\omega = 1$ rad/sec).

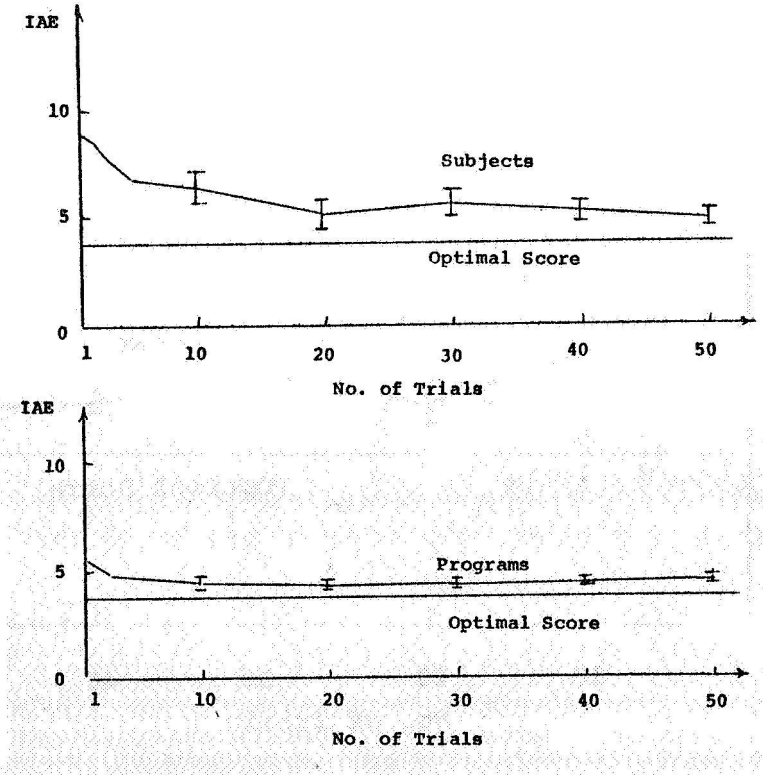


Figure 4.—Harmonic oscillator learning curves ($\omega = 1$ rad/sec).

REFERENCES

1. Preyss, A. E.; and Meiry, J. L.: Stochastic Modeling of Human Learning Behavior. Third Annual NASA-University Conference on Manual Control. SP-144, 1967.
2. McRuer, D. T.; Graham, D.; Krendel, E.; and Reisener, W. C.: Human Pilot Dynamics in Compensatory Systems—Theory, Models, and Experiments with Controlled Element and Forcing Function Variations. AFFDL-TR-65-15, July 1965.
3. Athans, M.; and Falb, P. L.: Optimal Control—An Introduction to the Theory and Its Applications. McGraw-Hill Book Co., Inc., 1966.
4. Mann, H. B.; and Whitney, D. R.: On a Test of Whether One of Two Random Variables is Stochastically Larger than the Other. Ann. Math. Stat., vol. 18, 1947, pp. 50-60.
5. Chien, T. T.: Human Learning Behavior in Manual Control Tasks. S.M. Thesis, MIT, Sept. 1967.